A model of the large-scale organization of chromatin

Mariano Barbieri*, Mita Chotalia†, James Fraser‡, Liron-Mark Lavitas†, Josée Dostie‡, Ana Pombo† and Mario Nicodemi*[1]

*Università di Napoli 'Federico II', Dipartimento di Scienze Fisiche, INFN Sezione di Napoli, CNR-SPIN, Complesso Universitario di Monte S. Angelo, Via Cintia, 80126 Napoli, Italy, †Genome Function Group, MRC Clinical Sciences Centre, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, U.K., and ‡Department of Biochemistry and Goodman Cancer Research Center, McGill University, 3655 Promenade Sir-William-Osler, Montréal, Québec, Canada, H3G 1Y6

Abstract

In the cell nucleus, chromosomes have a complex spatial organization, spanning several length scales, which serves vital functional purposes. It is unknown, however, how their three-dimensional architecture is orchestrated. In the present article, we review the application of a model based on classical polymer physics, the strings and binders switch model, to explain the molecular mechanisms of chromatin self-organization. We explore the scenario where chromatin architecture is shaped and regulated by the interactions of chromosomes with diffusing DNA-binding factors via thermodynamics mechanisms and compare it with available experimental data.

Introduction

It is now established that the chromosomes of higher eukaryotes have a complex spatial architecture within the cell nucleus, which serves vital functional purposes [1,2]. Such an organization is found at the scale of single genes which, for instance, contact their remote regulatory sequences, forming chromatin loops, or associate with transcription factories. Chromatin contacts span up to the scale of entire chromosomes, which are known to occupy specific territories in the nucleus. Interaction maps across specific chromosomes show preferential interactions, although it is still a mystery how such complex patterns are regulated and self-organized.

The development of new technologies, based on 3C (chromosome conformation capture) methods, has opened the experimental exploration of the contact maps of chromatin at genomic scales. In particular, Hi-C (high-throughput chromosome capture) technologies have revealed that the average contact probability, $P_c(s)$, of two loci having a genomic distance, s, scales as a power law [3], at least in the 0.5–7 Mb domain: $P_c(s) \sim 1/s^{\alpha}$. Such a discovery suggests a 'scale-free' organization in space. Although a power-law behaviour is expected in polymer models, the value of the exponent, $\alpha = 1.08$, reported in the first Hi-C study [3], is not found in usual equilibrium polymer systems (see [1,3], and references therein). Conversely, it was noticed [3] that a specific transient state exists (the fractal globule) of an ideal polymer chain (encountered, e.g., during decompaction) where the exponent of the contact probability is $\alpha = 1$, i.e. very close to the experimental one. The mean square distance between loci, $R(s)$, is also predicted to scale as a power law: $R(s) \sim s^{1/3}$. Yet, the hypothesis that chromatin conformations might be described by a single 'universal' state (transient or not) conflicts with the observation that a variety of different chromatin folding states exists, such as compact heterochromatin and open euchromatin.

In fact, we calculated $P_c(s)$ from genome-wide average data from a new wave of experiments and techniques [3–5], which shows that different cell types and systems have different exponents, α, spanning a range of values approximately from 0.8 to 1.6 (Figure 1). Furthermore, within a given cell system, different chromosomes are also characterized by contact probabilities with different exponents (Figure 1). Thus α does not capture a 'universal' property of chromatin folding and we have to explain the origin of the range of observed values.

The SBS (strings and binders switch) model

To describe the organization of chromatin inside the cell nucleus, we considered the scenario where chromosome architecture is shaped by the interactions of chromatin with the nuclear envelope, with nuclear bodies and, of course, DNA-binding molecular factors. In particular, for the sake of simplicity, in the present article, we focus on the interaction with diffusing molecules. We consider a polymer model, the SBS model where chromatin is represented as a SAW (self-avoiding walk) polymer chain with binding sites for diffusing molecules having a molar concentration c_m, and affinity E_x [6,7]. In this system, binding of diffusing molecules with more than one attachment site along the polymer can form bridges, i.e. chromatin contacts and loops (see snapshots in Figures 2 and 4).

The equilibrium and dynamics properties of such a system are subject to the laws of physics and can be studied by computer simulations. Within such a framework, we focus on questions such as: (i) how are different stable

Key words: chromatin organization, chromosome capture, polymer physics, self-organization, strings and binders switch model, thermodynamics.

Abbreviations used: FISH, fluorescence *in situ* hybridization; Hi-C, high-throughput chromosome capture; SAW, self-avoiding walk; SBS, strings and binders switch.

[1]To whom correspondence should be addressed (email mario.nicodemi@na.infn.it).

Figure 1 | **Chromatin in different cell systems has different contact probabilities and exponents**

Left: comparison between the average contact probability, $P_c(s)$, recorded in the different experiments and cell lines illustrated. Different behaviours and exponents are found in different systems. TCC, tethered conformation capture. Right: also within a given system, different chromosomes have a different $P_c(s)$ with different exponents (IMR90 Hi-C data from [5]). Typically, chromosome X is found above the genome-wide average, whereas chromosome 19 is below. Chr., chromosome.

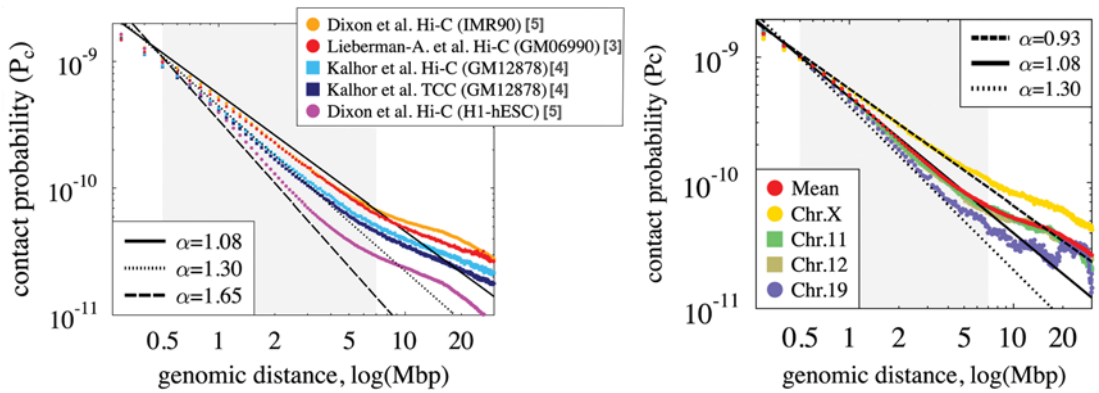
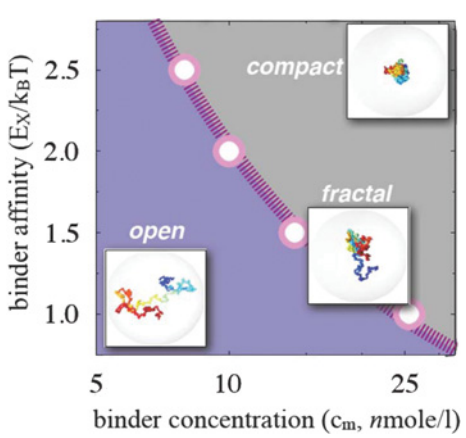


Figure 2 | **Phase diagram of the SBS model in the plane of its two control parameters, the concentration and affinity of binders, c_m and E_x**

The system is in an open and randomly folded conformation below the transition line, $C_{tr}(E_x)$ (broken curve), it folds in a compact conformation above it, and it takes the structure of the Θ-point state (which we dub 'fractal' here to clearly distinguish it from the others) at the transition line.



conformations (and their scaling properties) established?; (ii) how are architectural changes reliably regulated?; and (iii) is the scenario identified through the SBS model compatible with experimental Hi-C and FISH (fluorescence *in situ* hybridization) data?

The mechanisms of self-organization of the SBS model and its emerging stable states

The phase diagram of the SBS model is shown in Figure 2: a thermodynamic transition line separates its different phases which correspond to different stable architectural classes [8]. Below, at or above threshold, the polymer conformation corresponds to: (i) the open state, (ii) the Θ-point state (which we refer to below as the Θ-point fractal state to distinguish it from the other states), and (iii) the compact state. The energy and concentration scale predicted a fall in the range expected from biological considerations: the transition can be observed in the weak biochemical energy range and, correspondingly, at concentrations of the order of fractions of micromol/litre, typical of transcription factors in the nucleus. Importantly, in such a scenario, conformational changes can be regulated sharply and reliably (i.e. switch-like) by crossing the transition threshold (with no need for fine-tuning of the parameters) via simple strategies such as up- or down-regulation of the concentration of the polymer-binding molecules (i.e. by acting on c_m) or by chemical modifications of the binding sites along the polymer (i.e. acting on E_x), which correspond to well-described cell strategies. In summary, in the SBS model, the polymer-folding state depends on the concentration/affinity of the bridging molecules, and its architectural classes correspond to the three stable emergent states, whereas a variety of off-equilibrium conformations exists (including the fractal globule).

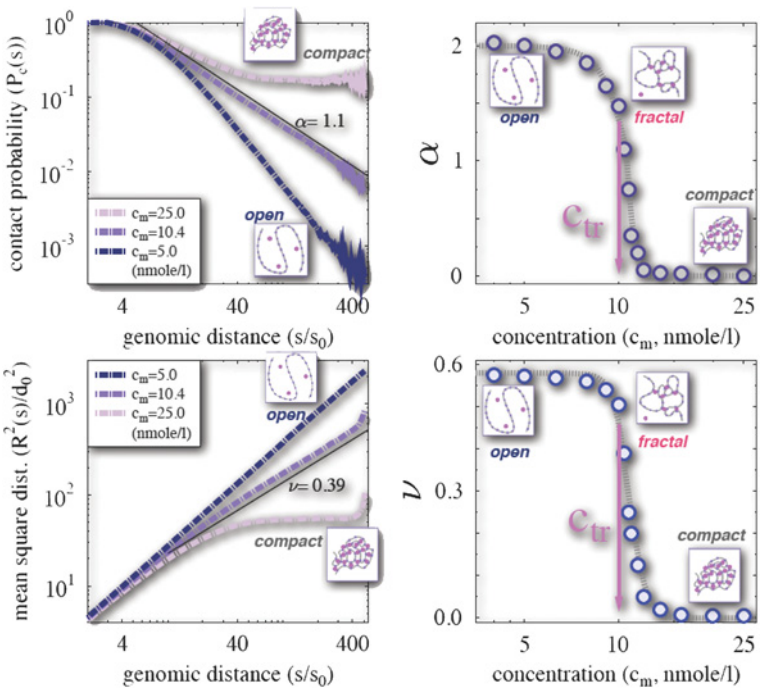
Comparison of the SBS model with experimental data

In the SBS model, the contact probability, $P_c(s)$, and the mean square distance, $R^2(s)$, are power laws at large genomic separations, s, with exponents depending on the concentration/affinity of binders (Figure 3): $P_c(s) \approx 1/s^{\alpha}$, with $\alpha = \alpha(c_m, E_x)$ and $R^2(s) \approx s^{2\nu}$, with $\nu = \nu(c_m, E_x)$. The value of the power law exponents does not change within a phase: it has a sigmoid shape as a function of c_m or E_x, with an inflection point at the phase transition.

In fact, chromatin is likely to be a mixture of differently folded genomic regions, as suggested for instance by FISH

Figure 3 | Contact probability and mean square distance in the SBS model

Top left: the average contact probability, $P_c(s)$, is shown for two sites having a contour distance, s, along the SBS model polymer. At large s, a power law is found, $P_c(s) \approx 1/s^{\alpha}$. Top right: the exponent α has a sigmoid behaviour as a function of c_m. Bottom left: also the mean square distance, $R^2(s)$, has a power law behaviour and exhibits a plateau in the 'compact' phase. Its exponent has a sigmoid shape too (bottom right).



experiments. To explore the important effects of variable folding on the contact probability, we considered a mixture of SBS model systems containing a fraction, f, of open polymers (where $\alpha = 2.1$; Figure 3) and a fraction, $1-f$, of compact polymers (where $\alpha = 0.0$). The exponent α of the corresponding average $P_c(s)$ depends on the fraction f (Figure 4, left-hand panel). A value $\alpha = 1.08$ can be found for $f \approx 0.60$, in a range of s that spans one order of magnitude, as observed in the Hi-C data. In cases where the fraction of open polymers is decreased down to $f = 0.45$, the mixture has $\alpha = 0.93$, in the same s range, a value close to the one found, for instance, for chromosome X in the human female cell lines GM06990 and IMR90 (compare Figure 1, left-hand panel with Figure 4, left-hand panel). Analogously, $f = 0.80$ gives $\alpha = 1.3$, as for chromosome 19 in the same cell lines. This simple exercise illustrates the fact that the values of α reported in Hi-C experiments reflect genome-wide averages of contact probabilities derived from a population of states, which disregard the variety of behaviours of different chromosomes or different genomic loci.

In summary, our analyses of Hi-C data show that the contact probability has different exponents in different chromosomes and systems, all of which fall in the range predicted by SBS model (0–2.1). Interestingly, the mean spatial distance of specific loci measured by single cells analyses by FISH are also fully described by the SBS model [7].

An investigation of the moment ratio $<R^4>/<R^2>^2$ can help to distinguish the different behaviours of the different loci. In polymer models, the moment ratio $<R^4>/<R^2>^2$ takes defined values, e.g. in the SAW polymer it is equal to 3/2. In the SBS model, $<R^4>/<R^2>^2$ depends on the concentration of binders: in the open and compact state, it is indeed 3/2, but it has a sharp peak in the transition region (the Θ-point 'fractal' region; Figure 5). Measurements of $<R^4>/<R^2>^2$ can therefore be used to probe whether a locus is open, Θ-point fractal or fully compacted. Importantly, FISH data [9] show that the moment ratio $<R^4>/<R^2>^2$ changes with genomic locus and genomic distance, along single chromosomes (Figure 5), spanning the very same range of values predicted by the SBS model. In fact, the SBS is currently the only model that explains the observed range of $<R^4>/<R^2>^2$. The experimental FISH results for a small number of regions suggest that chromatin is likely to be more than a mixture of open and closed states, as states typical of the transition region around the Θ-point appear to be common.

Discussion

The picture emerging from the available data and from the insights given by the SBS model (Figure 4, right-hand panel) is that chromatin is a complex mixture of differently folded regions. A specific locus in a given cell can be folded in different conformations. The SBS model describes some

Figure 4 | Contact probability in a mixture of states of the SBS model

Left: the average $P_c(s)$ is shown for mixtures of open and compact polymers in the SBS model: f is the fraction of open polymers and $1-f$ is the fraction of compact polymers in each mixture. $P_c(s)$ and its exponent α, in the highlighted decade, depend on f. Interestingly, by simply changing the fraction of open chromatin, the entire range of Hi-C exponents of Figure 1 is recovered. Right: pictorial view of the SBS scenario. Chromatin is a complex mixture of differently folded regions. A specific locus (in a given cell) can be folded in different conformations. The SBS model describes some of the mechanisms of self-organization, according to the affinity and concentration of its local binders. The Hi-C average contact probability, $P_c(s)$, and its exponent α, is an average over loci in different states, e.g. open, Θ-point 'fractal' and closed, and depends on relative fractions.

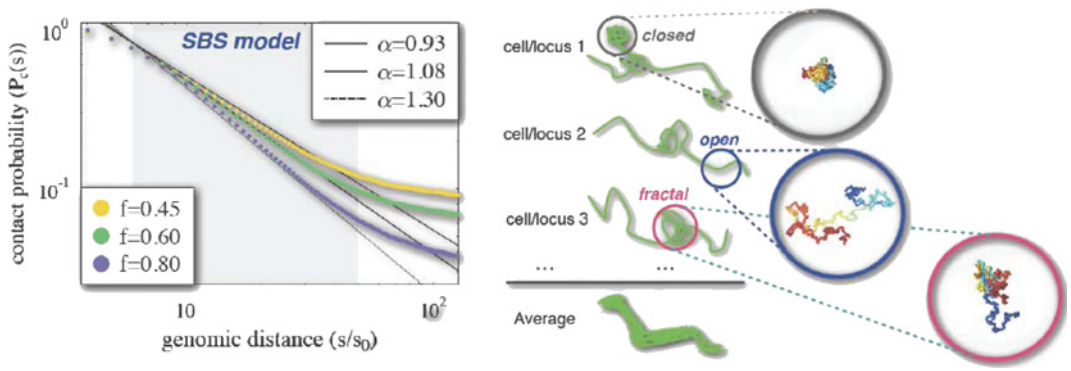
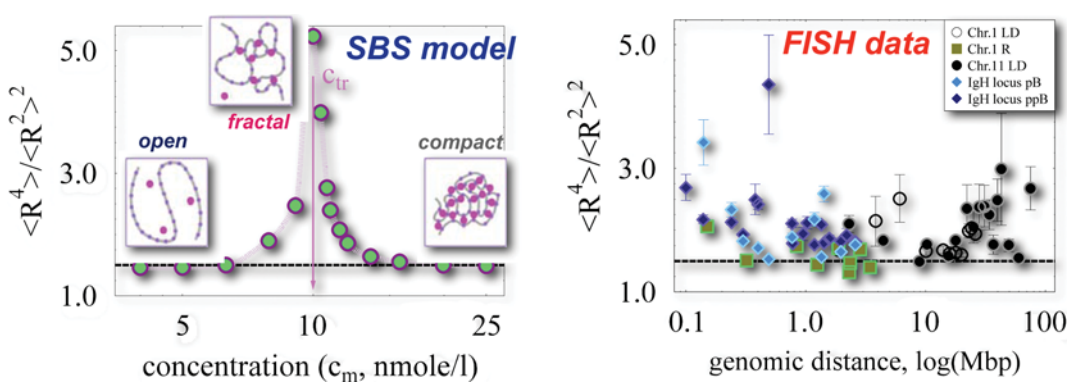


Figure 5 | The SBS model also explains the observed range of values of the moment ratio

Data suggest that the Θ-point state can be found in real cell chromatin. Left: in the SBS model, the moment ratio, $<R^4>/<R^2>^2$, depends on the concentration of binders. In both the open and compact states, it is 1.5, but it has a sharp peak at the transition region. It spans the same range found experimentally (right). Experimental data shown in the right-hand panel seem to suggest that chromatin can be found in both the open and compact state (where $<R^4>/<R^2>^2$ is 1.5) as well as around the Θ-point transition state (where $<R^4>/<R^2>^2$ is well above 1.5).



of the mechanisms of its self-organization. The locus has $\alpha \approx 0$, 1.5 or 2, if in a closed, Θ-point fractal or open state respectively, according to its local binders. The Hi-C $P_c(s)$, and its exponent $\alpha \approx 1$, represents an average over loci in different states, e.g. open, Θ-point fractal and closed, and depends on their relative fractions. We show that additional parameters, such as the moment ratio, $<R^4>/<R^2>^2$, can help to dissect the conformational states of chromatin.

In summary, the SBS model [6,7] illustrates how the architectural patterns of a polymer can be shaped and regulated by the interactions with randomly diffusing binding molecules via thermodynamic mechanisms. Folding classes correspond to stable emergent phases, although a variety of transient conformations can be described by the SBS model, including the fractal globule. Conformational changes can be sharply controlled by simple strategies, e.g. protein up-regulation or chromatin modification. In fact, the model can be also used to model specific chromosomal loci as the *Xist* locus [10–12]. Crowding and entanglement effects can be also considered as they can have important roles on the system dynamics as seen in other complex fluids [13–16]. Within the SBS scenario, the available FISH and Hi-C data, spanning from the average contact probability to the moment ratio of distances of loci, can be rationalized in a single framework.

The emerging picture is that chromatin is a complex mixture of differently folded regions, self-organized across spatial scales by SBS-like physical mechanisms.

References

1 Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G. and Cremer, T. (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat. Rev. Genet. **8**, 104–115
2 Pombo, A. and Branco, M.R. (2007) Functional organisation of the genome during interphase. Curr. Opin. Genet. Dev. **17**, 451–455
3 Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O. et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science **326**, 289–293
4 Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F. and Chen, L. (2011) Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat. Biotechnol. **30**, 90–98
5 Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature **485**, 376–380
6 Nicodemi, M. and Prisco, A. (2009) Thermodynamic pathways to genome spatial organization in the cell nucleus. Biophys. J. **96**, 2168–2177
7 Barbieri, M., Chotalia, M., Fraser, J., Lavitas, L.M., Dostie, J., Pombo, A. and Nicodemi, M. (2012) Complexity of chromatin folding is captured by the strings and binders switch model. Proc. Natl. Acad. Sci. U.S.A. **109**, 16173–1678
8 de Gennes, P.G. (1979) Scaling Concepts in Polymer Physics, Cornell University Press, Ithaca
9 Mateos-Langerak, J., Bohn, M., de Leeuw, W., Giromus, O., Manders, E.M.M., Verschure, P.J., Idemans, M.H.G., Gierman, H.J., Heermann, D.W., van Driel, R. and Goetze, S. (2009) Spatially confined folding of chromatin in the interphase nucleus. Proc. Natl. Acad. Sci. U.S.A. **106**, 3812–3817
10 Scialdone, A., Cataudella, I., Barbieri, M., Prisco, A. and Nicodemi, M. (2011) Conformation regulation of the X chromosome inactivation center: a model. PLoS Comp. Biol. **7**, e1002229
11 Nicodemi, M. and Prisco, A. (2007) Symmetry-breaking model for X-chromosome inactivation. Phys. Rev. Lett. **98**, 108104
12 Scialdone, A., Barbieri, M., Pallotti, D. and Nicodemi, M. (2011) Mean-field theory of the symmetry breaking model for X chromosome inactivation. Prog. Theor. Phys. Suppl. **191**, 40–45
13 Tarzia, M., de Candia, A., Fierro, A., Nicodemi, M. and Coniglio, A. (2004) Glass transition in granular media. Europhys. Lett. **66**, 531–537
14 Coniglio, A. and Nicodemi, M. (2000) The jamming transition of granular media. J. Phys.: Condens. Matter **12**, 6601–6610
15 Coniglio, A., Caglioti, E., Herrmann, H.J., Loreto, V. and Nicodemi, M. (1998) Segregation of granular mixtures in presence of compaction. Europhys. Lett. **43**, 591–597
16 Nicodemi, M. and Jensen, H.J. (2001) Creep of superconducting vortices in the limit of vanishing temperature: a finger-print of off-equilibrium dynamics. Phys. Rev. Lett. **86**, 4378–4381

Received 13 September 2012
doi:10.1042/BST20120238